

AD 653726

657-1-D090

DOUGLAS PAPER NO. 4496

**TECHNICAL WRITING:
A KEY TO COMPUTERIZED
INFORMATION RETRIEVAL**

PREPARED BY
S O JONES
CORPORATE LIBRARIAN
DOUGLAS AIRCRAFT COMPANY, INC.

PRESENTED TO
THE 14TH INTERNATIONAL TECHNICAL COMMUNICATIONS
CONFERENCE OF THE SOCIETY OF
TECHNICAL WRITERS AND PUBLISHERS
CHICAGO, ILLINOIS
24 - 27 MAY 1967



JUN 28 1967

DOUGLAS CORPORATE OFFICE

ARCHIVE COPY

F
24

Best Available Copy

TECHNICAL WRITING:
A KEY TO COMPUTERIZED INFORMATION RETRIEVAL

by

Shirli O. Jones
Corporate Librarian
Douglas Aircraft Company, Inc.

ABSTRACT

The future success of mechanized retrieval of information from published material lies, in a large part, in the hands of the technical writer. The context is only as usable as its format when current retrieval methods are used. Examples are given of confusion arising through nondefinitive titles, particularly where permuted indexes, such as Key-Word-In-Context (KWIC) and Key-Word-Out-of-Context (KWOC), are used.

The appearance of an article in several different publications is another source of difficulty for the information retrieval system. How the creation of separate entries for the same item can damage the integrity of an information system is shown.

A plea is made to the technical writers for understanding the needs of the second generation of reader. Steps which can be taken without censoring the language by replacing all synonyms and near-synonyms with a single term are indicated. The use of meaningful titles, standardization of citations, and several other areas are explored from the viewpoint of the information specialist.

FIGURES

1	KWIC and KWOC Indexes	4
2	Title Repetition by Word	5
3	KWOC Index With Terms Repeated	6
4	List of Words Prevented From Indexing	7
5	Effects of Hyphen Usage Upon Computer Sort	8
6	Effects of Singular-Plural Usage Upon Computer Sort	9
7	Index Card With Multiple Corporate Authors	12
8	Variations of Single Corporate Author	14
9	Automatic Index	16

Preceding Page Blank

INTRODUCTION

Each year, more and more persons decry the proliferation of the printed word. Available statistics prove that unless restraint is applied, mankind will be submerged in published information by 1975. Concern over how to obtain essential information from this vast amount of material results in a rash of papers such as this. Information specialists involved with the retrieval of necessary data bemoan the unnecessary generation of papers, but of more importance is more careful consideration in both the preparation and the publication of all papers. Such a specialist is not in a position to determine the relative value of one document in comparison with another, but he is in a position to help establish guidelines suitable for advanced computer techniques that will provide the keys to necessary storage and retrieval capabilities for the future. Although information retrieval systems covering all degrees of complexity have been introduced, none can produce a product better than the original material. Therefore, authors and technical writers must cooperate with the information specialist to ensure the rapid and effective retrieval of valuable data.

Information retrieval, in general, relies on two forms of index--subject headings and concept coordination. The more familiar form of the subject heading index is the public library card. The concept coordination system uses "uniform terms" or "descriptors"--a single term or concept for every major topic in a document. These terms can then be coordinated with other terms or concepts for flexible, specialized retrieval. Bibliographic information, such as personal author, corporate author, and title, is handled in much the same way, regardless of the type of retrieval system used. The equipment used in creating and manipulating the index can range all the way from a typist typing on 3 x 5 file cards to a fully computerized system for storage and automated retrieval in various formats. In every case, the technical writer's importance is evident, but it is never so obvious as in the computerized storage and retrieval system. Thus, the problems associated with this area require special emphasis.

DISCUSSION

Since 1959, when Douglas actively began to design its mechanized information retrieval program, many different systems have been developed. Government agencies, such as NASA's Scientific & Technical Information Division and DOD's Defense Documentation Center, have built massive systems utilizing the most sophisticated of computer equipment. Several private industries have expended vast quantities of time and money to provide their personnel with retrieval capabilities. Organizations such as the American Society for Metals and the American Chemical Society have called upon their professional talents to produce systems designed to cope with the complexities of their scientific disciplines. These are all major efforts, but the needs of the smaller systems are of equal importance. And either type of system must face much the same problems. Some of these problems, which will be described in detail, are the responsibility of the technical writers and/or publishers.

One of the most common forms of mechanized index is the Key-Word-In-Context or KWIC index, and a variation called KWOC, or Key-Word-Out-of-Context (Figure 1). As can be seen, these indexes consist of a sorting of the title with each word repeated in alphabetical order. The title will appear as many times as there are significant words in the title (Figure 2). The KWOC index provides a more readable format (Figure 3) but follows the same principle. To prevent listings of insignificant terms (such as "the" and "of"), a stop list is input to the computer. Chemical Titles employs a stop list of 1,335 "insignificant" words (Figure 4). Examination of this list indicates that titles such as "Laboratory Systems & Procedures" would be completely eliminated.

Another problem occurs with the inconsistent use of hyphens. In Figure 5, there are several interesting entries. Under "low," the first inconsistency occurs with "low energy"--two entries are unhyphenated, three hyphenated. Next is "low molecular weight"--two entries without hyphens, one with.

Figure 1. KWIC and KWOC Indexes

IC LESIONS IN THE BROAD USE/	BREASTED BRONZE TURKEY/ THE INFLUENC	7310
OF EXERCISE AND O ₂ % ON CREATINE TREATMENT UPON	BREATH HOLDING AT BEGINNING OF EXERC	5742
KIN OF THE COMMON SOLE/	BREATH HOLDING/ INFLUENCE	5740
GITAL PRESSURE AND DEEP	BREATHING ENERGETIC OUTPUT AND AEROB	4984
EFFECT OF O ₂ %	BREATHING MOVEMENTS IN ENTOSCELLA-SO	8326
SYSTEMIC EFFECTS AFTER	BREATHING% ON PATIENTS WITH ADAMS STO	5522
URING POSITIVE PRESSURE	BREATHING ON PULMONARY COMPLIANCE/	5745
P AS RELATED TO SEASON,	BREATHING POTENTI MEDICATED AEROSOLS/	5754
THE DIVERSION AND THE	BREATHING/ DIAPHRAGM ACTIVITY AND TH	5734
LEAST TERM/ THE	BREED, SEX, AND SEMEN QUALITY/ THE T	7254
CATIONSOF THE CHOICE OF	BREEDING BEHAVIOR OF THE RAVEN WITH	8544
SE FREE PATHOGEN FREE	BREEDING BIOLOGY AND ETHOLOGY OF THE	4595
LICATION OF GENETICS TO	BREEDING BIOTYPE/ THE STATE OF DEVELO	4593
CAL PROBLEMS IN LUCERNE	BREEDING COLONY/ PROGRESS REPORT. DI	7371
E HYBRID FORAGE SORGHUM	BREEDING DOMESTIC ANIMALS/ APP	6566
FLORIDA/ AUTUMNAL	BRECUING IN CONNECTION WITH FERTILIT	7908
TAMIC ACID FORMATION IN	BREEDING IN HUNGARY/ THE PRESENT POS	7907
CAL COMPOSITION OF SAKE	BREEDING OF BOAT-TAILED GRACKLES IN	8553
OUPING/ STUDIES ON SAKE	BREVIBACTERIUM-FLAVUM/ SIGNIFICANCE	6697
WATER/ STUDIES ON SAKE	BREWING WATER AND THEIR GROUPING/ ST	5159
MONG COMPONENTS OF SAKE	BREWING WATER. CHEMICAL COMPOSITION	5159
CK BY ELECTRONIC MEANS/	BREWING WATER. RELATIONSHIP AMONG CO	5160
THE CURING PROCESS FOR	BREWING WATER/ STUDIES ON SAKE BREWI	5160
E RESPONSES IN A SIMPLE	BRIDGING OF INTERRUPTED ATRIO VENTRI	5436
A DEVICE FOR FEEDING	BRIGHT-LEAF TOBACCO/ STEADY-STATE TH	7848
OF PLANT PHYSIOLOGISTS,	BRIGHTNESS DISCRIMINATION UNDER DIFF	4525
OF PLANT PHYSIOLOGISTS,	BRINE SHRIMP TO FISHES/	4685
OF PLANT PHYSIOLOGISTS,	BRISBANE MAY 1961/ LEAF TEMPERATURE	7716
THE WITH HIGH MOUNTAINS	BRISBANE MAY 1961/ THE CUTICLE IN EU	7697
THE GENUS ASELLUS IN	BRISBANE MAY 1961/ THE EFFECTS OF MO	7703
CIES OF LITHOSID NEW TO	BRISKETD DISEASE AND IN EXPERIMENTAL	7274
N POTATO CROPS IN GREAT	BRITAIN/	8362
MAN-MADE ACTIVITIES IN	BRITAIN/ PELOSI-A-OBUSA-HERRICH-SCHA	8419
IZATION IN THE HANDS OF	BRITAIN, 1952-60/ EXTENT OF PROTECTI	8199
IN THE WEST INDIES AND	BRITISH COLUMBIA/ THE EFFECTS ON FRE	4643
BIOLOGICAL FLORA OF THE	BRITISH FISH FILLETTERS/ COLD VASODIL	6334
BIOLOGICAL FLORA OF THE	BRITISH GUIANA/ EASTERN EQUINE ENCEP	7487
ABERRATIONS OF	BRITISH ISLES. ARRHENATHERUM-ELATIUS	7682
INDIES, THE GUAJANAS, AND	BRITISH ISLES. FRAXINUS-EXCELSIOR-L/	7690
VITIUS-EDULIS-L. IN THE	BRITISH LEPIDOPTERA/	8413
CLEROTIC LESIONS IN THE	BRITISH-MONOMORAS/ PRESENT AND POTENT	7850
OF THE DISTRIBUTION OF	BRITISH-ISLES AND THEIR RELATIONSHIP	8448
F ISOLATED FRACTION OF	BROAD BREASTED BRONZE TURKEY/ THE IN	7310
RAIN DIFFERENTIATION OF	BROAD-LEAVED EVERGREEN TREES, BASED	4564
TWO VARIETIES OF SMOOTH	BROKEN CHLOROPLASTS/ HILL ACTIVITY O	7728
OL STUDIES OF SOME WEED	BROMEGRASS MOSAIC VIRUS/ PURIFICATIO	8157
A PROTEO LYtic ENZYME,	BROMEGRASS/ EFFECT OF CERTAIN FERTIL	7916
SPECIAL REFERENCE TO 3-	BROMEGRASSES/ LIFE CYCLES AND CONTR	7958
	BROMELAIN/ SYSTEMIC BIO CHEMICAL CHA	6259
	MONO DEOXY URIDYL/ IONIZATION OF D	4284

Figure 2. Title Repetition by Word

NON-IRRADIATED	ABSORPTION OF D-GALACTOSE BY SEGMENTS OF INTESTI- NE FROM ACTIVE AND HIBERNATING, IRRADIATED AND NON-IRRADIATED GROUND SQUIRRELS; CITELLUS TAI- DEOCMINUTUS. NASA 903-210021(R) 82-60 0276	NUCLEAR	ETIC BLACKOUT FOLLOWING A HIGH ALTITUDE NUCLEAR EXPLOSION. 80-291 151(R) 82-60 0372
NON-ISOTHERMAL	CORRELATIONS IN A NON-ISOTHERMAL PLASMA. AD-290 0277(R) 82-60 0278	NUCLEAR	ACCURATE NUCLEAR FUEL BURNUP ANALYSIS. GR-9021(R) 81-10 0362
NON-LINEAR	INVESTIGATION OF MICROWAVE NON-LINEAR EFFECTS UTILIZING PERFORAMIC MATERIALS. AD-290 0278(R) 82-60 0287	NUCLEAR	APPLICATION OF NUCLEAR POWER SUPPLIES TO SPACE SYSTEMS. 80-291 152(R) 82-60 0371
NON-METALLIC	BIBLIOGRAPHY AND TABULATION OF CAMPING PROPERTY TESTS OF NON-METALLIC MATERIALS. AD-290 0279(R) 82-60 0292	NUCLEAR	CALIFORNIA-VIRGINIA NUCLEAR POWER ASSOCIATES, I NCL RESEARCH AND DEVELOPMENT PROGRAM QUARTER Y PROGRESS REPORT FOR THE PERIOD APRIL - JUNE 1962. CYAN-151(R) 82-60 0379
NON-MILITARY	NOTES ON NON-MILITARY MEASURES IN CONTROL OF INSURGENCY. AD-290 237(R) 81-60 0296	NUCLEAR	COMPUTER PROGRAMS FOR OPTIMUM START-UP OF NUCL EAR PROPULSION SYSTEMS. ED-14730(R) 81-10 0272
NON-MOVING	JUDGEMENTS OF VISUAL VELOCITY AS A FUNCTION OF THE LENGTH OF OBSERVATION TIME OF MOVING OR NO M-MOVING STIMULUS. PB 162 546(R) 81-60 0295	NUCLEAR	DOSE-TIME-DISTANCE CURVES FOR CLOSE-IN FALLOUT FOR LOW YIELD LAND-SURFACE NUCLEAR DETONATION S. PB 162 561(R) 81-60 0373
NON-RELATIVISTIC	TABLES OF NON-RELATIVISTIC ELECTRON TRAJECTO RIES FOR FIELD EMISSION ON METALS. AD-290 0280(R) 82-60 0299	NUCLEAR	EXTENDED CERAMIC NUCLEAR FUEL DEVELOPMENT PROG RAM. ACNP-62530(R) 82-60 0392
NON-SIMILAR	NON-SIMILAR NUMERICAL METHODS OF SOLUTION FOR ELECTRODE BOUNDARY LAYERS IN A CROSSED FIELD A CCCELERATOR. AD-290 929(R) 82-60 0393	NUCLEAR	FEASIBILITY DETERMINATION OF A NUCLEAR THERMI CIC SPACE POWER PLANT. AD-290 0281(R) 82-60 0393
NONDESTRUCTIVE	NONDESTRUCTIVE SYSTEM FOR INSPECTION OF GR/REIN FORCED GLASS-REINFORCED PLASTIC (GR/GRP) CASES. AD-290 0282(R) 82-60 0394	NUCLEAR	HIGH-ENERGY NUCLEAR PHYSICS RESEARCH PROGRAM AD-291 140(R) 81-60 0376
NONDESTRUCTIVE	X-RAY IMAGE SYSTEM FOR NONDESTRUCTIVE TESTING OF SOLID PAULLENT MISSILE CASE WALLS AND WEL DMENTS. AD-290 231(R) 82-60 0437	NUCLEAR	HIGH-INTENSITY NUCLEAR REACTIONS OF MEDIUM WITH INCIDENT PROTONS AND HELIUM IONS. AD-290 0283(R) 82-25 0372
NONDISSIPATIVE	MAGNETOHYDRODYNAMIC STABILITY OF VORTEX FLOW - A NONDISPERSIVE, INCOMPRESSIBLE ANALYSIS. AD-290 0284(R) 82-60 0239	NUCLEAR	INVESTIGATIONS ON THE DIRECT CONVERSION OF NUC LEAR FISSION ENERGY TO ELECTRICAL ENERGY IN A PLASMA DIODE. AD-290 727(R) 82-60 0385
NONEQUILIBRIUM	SCALE EFFECTS FOR NONEQUILIBRIUM CONDUCTIVE HE AT TRANSFER WITH SIMULTANEOUS GAS-PHASE AND SU RFACE CHEMICAL REACTIONS. APPLICATION TO HYPER SONIC FLIGHT AT HIGH ALTITUDES. AD-291 032(R) 81-60 0225	NUCLEAR	NUCLEAR SUPERHEAT DEVELOPMENT PROGRAM. GRNC-254(R) 81-60 0386
NONLINEAR	APPLICATION OF VARIATIONAL EQUATION OF MOTION TO THE NONLINEAR VIBRATION ANALYSIS OF HOMOGEN EOUS AND LAYERED PLATES AND SHELLS. AD-290 0285(R) 82-60 0407	NUCLEAR	PRODUCTION OF TRITIUM BY CON-INDUCED NUCLEAR EXP LOSIONS IN SALTS. I. LABORATORY STUDIES OF ISOT OPIC EXCHANGE OF TRITIUM IN THE HYDROGEN-METAL SYSTEM. ORNL-333(R) 8-30 0417
NONLINEAR	EXTENSIONS IN THE SYNTHESIS OF TIME OPTIMAL OR BANG-BANG NONLINEAR CONTROL SYSTEMS. PART II. THE SYNTHESIS OF QUASI-STATIONARY OPTIMUM NON LINEAR CONTROL SYSTEMS. PB 162 147 547(R) 84-60 0285	NULL-ZONE	STRIKING EFFECT OF NUCLEAR EXPLOSION. AD-290 826(R) 821-00 0083
NONLINEAR	EXTENSIONS IN THE SYNTHESIS OF TIME OPTIMAL OR BANG-BANG NONLINEAR CONTROL SYSTEMS. PART II. THE SYNTHESIS OF QUASI-STATIONARY OPTIMUM NON LINEAR CONTROL SYSTEMS. PB 162 147 547(R) 84-60 0285	NUMBERS	THE NUCLEAR PROPERTIES OF RHENIUM. AD-291 180(R) 81-60 0310
NONLINEAR	NONLINEAR FLEXURAL VIBRATIONS OF SANDWICH PLAT ES. AD-290 0286(R) 82-60 0299	NUCLEAR	VARIATIONS IN THE TOTAL RADIATION DOSE RATE IN TH E THERMOSPHERE AFTER THE HIGH ALTITUDE NUCLEAR EXPLOSION. NASA 903-10464(R) 81-10 0142
NONLINEAR	OPTIMUM NONLINEAR CONTROL FOR ARBITRARY DISTUR BANCES. NASA 942-1980(R) 82-60 0482	NUCLEAR	4304 MARITIME NUCLEAR STEAM GENERATOR. GEMP-160(R) 80-10 0349
NONACURRENT	A TECHNIQUE FOR WARM-BAND TELEMETRY OF NORME CURRENT PULSES. AD-290 0287(R) 82-60 0577	NUCLEAR	THE ESTIMATION PROBLEM IN NULL-ZONE RECEP TION FEEDBACK SYSTEMS. AD-290 325(R) 81-60 0599
NONUNIFORM	ELECTROMAGNETIC SCATTERING FROM A SPHERICAL, NO NUNIFORM REGION. PART II. THE RADAR CROSS SEC TION. I. FOR A FEW POINTS. AD-290 614(R) 82-60 0747	NUCLEAR	FOCUSING OF A LASER BEAM INTO A SPHERICAL, NO NUNIFORM REGION. PART II. THE RADAR CROSS SEC TION. I. FOR A FEW POINTS. AD-290 614(R) 82-60 0747
NONUNIFORM	NON-ELECTROMAGNETIC SCATTERING FROM AN ASYMETRICAL NON UNIFORM MEDIUM. PART I. GENERAL. AD-290 614(R) 82-60 0748	NUCLEAR	FOCUSING OF A LASER BEAM INTO A SPHERICAL, NO NUNIFORM REGION. PART II. THE RADAR CROSS SEC TION. I. FOR A FEW POINTS. AD-290 614(R) 82-60 0748
NORMAL	PROBABILITY INTEGRALS OF MULTIVARIATE NORMAL A NORMAL	NUSTAGMUS	MANIPULATION OF ARCSUAL AND ITS EFFECTS ON HUM AN VESTIBULAR NYSTAGMUS INDUCED BY CALORIC IRR IGATION. AD-290 346(R) 81-60 0292
NORMAL	PROBABILITY INTEGRALS OF MULTIVARIATE NORMAL A NORMAL	OAK	A SAFETY REVIEW OF THE NUCLEAR POWER PLANT RIMINI FACILITY. OAK-TR-347(R) 82-60 0412
NORMS	NORMS FOR ARTIFICIAL LIGHTING. AD-290 935(R) 81-10 0734	OBJECTS	DRAG OF OBJECTS IN PARTICLE - LADEN AIR FLOW P HASE. IV. BLUNT BODIES AND COMPRESSIBILITY EFFE CTS. AD-291 170(R) 81-60 0792
NORTH	FACTORS INFLUENCING VASCULAR PLANT ZONATION IN NORTH CAROLINA SALTMARSHES. AD-290 938(R) 81-60 0803	OBSERVATORY	TUNTO FOREST SEISMOLOGICAL OBSERVATORY. AD-290 169(R) 81-60 0611
NORTH	SONAR STUDIES OF DEEP SETTLING LAGS IN THE NORTH PACIFIC. PB 162 427(R) 82-60 0578	OCEAN	A SAMPLE TEST EXPOSURE TO BACTERIA CORROSION AN D FOULING OF EQUIPMENT INSTALLED IN THE OCEAN. O CEAN. PB 162 428(R) 82-60 0595
NORTH	THE DEVELOPMENT OF RESCUE AND SURVIVAL TECHNI QUES IN THE NORTH AMERICAN ARCTIC. PB 162 940(R) 81-60 0805	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHURCH SEAS, SUMMER 1969. PART I. THE BERING STUDIES. PB 162 429(R) 82-60 0596
NOSE	THE FLORA OF HEALTHY DOGS. I. BACTERIA AND FUN GI OF THE NOSE, THROAT, AND LOWER RESPIRATORY TRACT. AD-290 939(R) 82-60 0578	OCEANOGRAPHIC	OCEANOGRAPHIC AND UNDERWATER ACOUSTICS RESE ARCH. VOL. 1. PB 162 430(R) 82-60 0597
NOZZLE	FABRICATION OF PYROLYtic GRAPHITE ROCKET NOZ ZLE COMPONENTS. PB 162 37(R) 81-10 0391	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHURCH SEAS, SUMMER 1969. PART IV. PHYSICAL OCEANOGR APIC STUDIES. VOL. 1. DESCRIPTIVE REPORT. PB 162 430-1(R) 82-60 0598
NOZZLE	FABRICATION OF PYROLYtic GRAPHITE ROCKET NOZ ZLE COMPONENTS. PB 162 37(R) 81-10 0393	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHURCH SEAS, SUMMER 1969. PART IV. PHYSICAL OCEANOGR APIC STUDIES. VOL. 1. DESCRIPTIVE REPORT. PB 162 430-1(R) 82-60 0598
NOZZLE	FABRICATION OF PYROLYtic GRAPHITE ROCKET NOZ ZLE COMPONENTS. PB 162 37(R) 81-10 0393	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHURCH SEAS, SUMMER 1969. PART IV. PHYSICAL OCEANOGR APIC STUDIES. VOL. 1. DESCRIPTIVE REPORT. PB 162 430-1(R) 82-60 0598
NOZZLES	HEAT TRANSFER IN PARTICULATE REACTOR IN SOL ID-FUELED MATERIALS. AD-290 354(R) 81-60 0280	OCEANOGRAPHIC	OCEANOGRAPHIC CRUISE TO THE BERING AND CHURCH SEAS, SUMMER 1969. PART IV. PHYSICAL OCEANOGR APIC STUDIES. VOL. 1. DESCRIPTIVE REPORT. PB 162 430-2(R) 82-60 0599
NOTE	DEVELOPMENT AND STANDARDIZATION OF FORMS 3 AND 4 OF THE NMC CONTRACT. STUDENT SELECTION TES T. AD-290 704(R) 81-10 0291	OCEANOGRAPHIC	PROCEEDINGS OF INTERINDUSTRIAL OCEANOGRAPHIC S YMPUSUM (IND. 1), BURBANK, CALIFORNIA, 5 JUNE 1962. PB 162 987(R) 82-60 0491
NOTC	EVALUATION OF RADIO AVIATION INDOCTRINATION F IELD TOURS FOR 1961-1962. AD-290 356(R) 81-60 0581	OCTYL	RUBBER ELASTICITY IN HIGHLY CROSSLINKED SYSTEM
NUCLEAR	A 7090 CODE FOR THE CALCULATION OF ELECTROPAK		

Figure 3. KWOC Index with Terms Repeated

Figure 4. List of Words Prevented from Indexing

COINE-131 IN BITCHES ON LOW AND MODERATE FAT INTAKE. ⁺	I	PSB-0111-0674
YIELDS. ⁺ EFFECTS OF LOW DOSES OF GAMMA RADIATION ON PLATE	I	JAR-0013-0487
ESIUM-84 (P,P/GAMMA) AT LOW ENERGIES. ⁺ MECHANISM IN MAGN	P	PHV-0126-2261
INELASTIC SCATTERING OF LOW ENERGY PHOTONS FROM NEON-20. ⁺	M	MAM-0039-0686
RUS-31 (P,GAMMA) SULFUR LOW ENERGY RESONANCES IN THE PHOSPHO	H	HSH-0037-0631
GLANDS. ⁺ EXTRACTION OF LOW MOLECULAR WEIGHT CORTICO TROPHIN	A	ACB-0042-0669
ID FROM C. ⁺ ISOLATION OF LOW MOLECULAR WEIGHT RIBO NUCLEIC AC	S	SCIN-0020-0488
RIVIZATION OF ETHYLENE AT LOW PRESSURES. ⁺ METHOD FOR POLYME	C	ANCE-0074-0977
CORTEX AS A CAUSE OF LOW RESISTANCE OF THE NEWBORN AND	P	PEST-00-06-088
IN THE + INFLUENCE OF A LOW SODIUM DIET ON DYNAMIC CHANGES I	V	VPIT-21-06-017
MELINES WITH CAUSTIC + LOW TEMPERATURE AGGLOMERATION OF HEP	V	IVUT-00-06-003
AND + OXIDES FROM THE LOW TEMPERATURE OXIDATION OF MAGN	J	JCOM-0004-0633
EFORMED ALUMINUM NATURE OF LOW TEMPERATURE TRANSFORMATIONS IN O	F	FMNT-0014-0750
OF LINEAR POLYMERS AT LOW TEMPERATURES, ⁺ VIBRATION SPECTRUM	D	DAMK-0147-0380
ETIC BALANCE FOR USE AT LOW TEMPERATURES. ⁺ GOUT MAGN	S	BCSJ-0035-2049
ON ADMIXTURE NUCLEI AT LOW TEMPERATURES. ⁺ MOSSBAUER EFFECT	D	DAMK-0147-0584
MUATION IN MAGNESIUM AT LOW TEMPERATURES. ⁺ ULTRA SONIC ATTE	J	JMPS-0017-1881
DIOSLYSIS OF N-METHANE AT LOW TEMPERATURES. ⁺ REACTIONS IN RA	C	JCP-0037-2496
NITRIDE INCLUSIONS IN LOW-CARBON HIGH-CHROMIUM STEEL. ⁺	I	IANN-02-06-073
MANGANESE REFINED FROM LOW-CONCENTRATION ORES CONTAINING	A	ARRN-0007-0361
IMPITATES ESTIMATION OF LOW-DENSITY LIPID PROTEIN IMMUNO PREC	C	CLCH-0005-0616
OF NEGATIVE HYDROGEN + LOW-ENERGY COLLISION CROSS SECTIONS	J	JCP-0037-2571
LEVELS OF MOLECULES BY LOW-ENERGY ELECTRON IMPACT SPECTROSC	J	JCP-0037-2497
ENERGY DISTRIBUTION IN LOW-ENERGY ELECTRON-PHOTON SHOWERS	P	PHV-0126-2352
ECTRUM OF INDIUM. ⁺ LOW-FREQUENCY LATTICE VIBRATIONAL SP	J	JCP-0037-2727
HERONIC CONVERTER WITH + LOW-FREQUENCY OSCILLATIONS IN A THER	C	CRAB-0016-0889
GANIC COMPOUNDS USING A LOW-LEVEL NEUTRON SOURCE. ⁺ IN OR	S	BCSJ-0035-2045
DEUTEROLYSIS RATES OF LOW-MOLECULAR-WEIGHT ALKANE SULFONYL	Z	ZPCL-0221-0103
HYDRIDE. ⁺ POSSIBLE LOW-PRESSURE POLYMORPHISM IN LITHIUM	J	JCP-0037-2730
OF TULENOVO OIL AFTER LOW-TEMPERATURE CATALYTIC TREATMENT.	C	CRAB-0016-0949
WITH DI CHLORIDES FOR LOW-TEMPERATURE DEWAXING OF OILS. ⁺	K	KTTH-07-12-011
LOW-TEMPERATURE POLYMERIZATION OF ET	A	AMCE-0074-0935

Figure 5. Effects of Hyphen Usage Upon Computer Sort

One of each appears under "low pressure." Following this is one of the most common: "low temperature," with eight unhyphenated entries to three hyphenated--an overwhelming verdict for abandoning the hyphen in this case.

An additional factor for consideration is that most computer sorting treats a hyphenated word as one word--that is, where "low temperature" appears as two separate words, the citation will appear under both "low" and "temperature." Where the hyphen is used, the citation will appear under "low," but not under "temperature."

Such details as the use of singular and plural word forms also can cause difficulty (Figure 6). "Line" and "lines" have been separated by seven entries of "linear" and "linearity." This may seem insignificant in the example, as shown, and it is, because the eye can scan enough of the page at one glance

FERTILIZATION. LINE AND P PLACEMENT EFFECTS ON REACTIONS IN HEATED LINE-ALUMINA MIXTURES. ⁻	SSRA-0026-0574
EFFECTS OF FERTILIZERS. LINE AND CULTIVATIONS ON YIELD.	JACH-0012-0535
ATION IN THE PROBLEMS, LIMITATIONS, AND THE FUTURE OF AUTOMETRY. ⁻	JBBC-0013-0321
DETECTION LIMITS IN RADIATION AND OPTICAL PYROMETERY WITHIN THE TEMPERATURE LIMITS 20-60-DEG. ⁻	ANYA-0162-0171A
LEAR MAGNETIC RESONANCE LINE OF ZEOLITIC WATER. ⁻	JOSA-0052-1307
MAGNETIC RESONANCE WIDE LINE SHAPES GENERATED BY TWO BROADEN	PLMS-62-11-053
MAGNETIC RESONANCE WIDE LINE SPECTRA. ⁻	CORE-0258-3400
N IODIDE. ⁻	ACSA-0016-2212
LINE STRENGTHS AND WIDTHS IN HYDROGEN	ACSA-0016-2149
MAGNETIC RESONANCE LINE WIDTH IN GARNET AND SPINEL TYPE	JCPB-0037-2699
ORGANIC ELECTROLYTES IN LINEAR AND CIRCULAR CHROMATOGRAPHY	FTVT-0064-3684
N-RANDOM DEGRADATION OF LINEAR CHAIN MOLECULES. ⁻	CANV-0007-0909
N-RANDOM DEGRADATION ON LINEAR CHAIN MOLECULES. ⁻	JUPS-0017-1614
ON THE ANTI FERRO MAGNETIC LINEAR CHAIN. ⁻	JUPS-0017-1696
SPECTRUM OF DIS-ORDERED THEORY FOR LINEAR COLLOIDS. ⁻	PMRV-0128-2131
OF THE COEFFICIENT OF LINEAR EXPANSION OF GLASSY PLASTICS.	JCPB-0037-2323
OF THE SPECIFIC HEAT OF LINEAR POLYMERS AT LOW TEMPERATURES.	PLMS-62-11-036
INVESTIGATION ON THE LINEARITY OF THE TEST CURVE FOR THE	DANK-0147-0500
CHANGE OF K(3S)- LINES AND THE VALUE OF THE INITIAL	ZACP-0192-0378
RY OF SPECTRAL HYDROGEN LINES IN A PLASMA. ⁻	PNMT-0014-0600
ASYMMET	DANK-0147-0342
EUROPIUM(II) EMISSION LINES IN EUROPIUM DI BENZOYL METHANE	JCPB-0037-2333
SITES OF THE MANGANESE LINES. ⁻	MGKF-0060-0623
OF THE DIFFERENCE OF THE BEN	DANK-0147-0636
MMA-COJUGATE SYSTEM OF LINOLEIC. ⁻	CHAL-0044-0463
COMPOUNDS WITH A CLOSED CYANATE IN ETHYLENIC LINKAGES. ANALYTICAL APPLICATIONS.	JACH-0230-0445
INHIBITION OF ENERGY- LINKED DI PHOSPHO PYRIDINE NUCLEOTID	DANK-0147-0106
ODD NUMBER OF MONOMERIC LINKS, FORMED IN THE THERMAL POLYMER	JACD-0030-0517
DUCT+ISOLATION OF PURE LINOLEATE AS ITS MERCURIC ACETATE A	IJAC-0025-0083
EUREA FRACTIONATION OF LINSEED OIL FATTY ACIDS. COMPARATIV	IGSB-0064-0106
CTIVITY OF LIPO PROTEIN LIPASE IN VARIOUS TISSUE SLICES. ⁻	AIPF-0140-0577
A GREATER ACTION OF PHOSPHO LIPASE ON A MAST. CELLS. ⁻	IGSB-0064-0105
TR	NNIK-0030-1004
PASE IN + LIPO PROTEIN LIPASE. ACTIVITY OF LIPO PROTEIN LI	NATU-0197-0070
YCERIDES BY PANCREATIC LIPASE. ⁻	UIJO-0066-0360
+ ENZYMIC HYDROLYSIS OF G	
T INTESTINAL + PHOSPHO LIPID COMPOSITION AND TURNOVER IN RA	
LIPID COMPOSITION OF TUMOR CELLS. ⁻	

Figure 6. Effects of Singular-Plural Usage Upon Computer Sort

to see both entries even though they are separated. If this index covered 5,000 to 6,000 items, these terms could appear so far apart that the searcher would not see them both, and thus would miss a portion of the indexed information.

Many of the titles used today have meaningful terminology because the authors were aware that the titles were likely to appear in KWIC or KWOC index formats. But what becomes of a title such "A Look Ahead," or "Structures and Economics," or "Pilot Studies and System Components Evaluation"? Each of these papers is concerned with means of handling information, and each was presented at a national symposium on Engineering Information, the theme of which was "A Coordinated Engineering Information System." The difficulty

in indexing titles such as these shows that even information specialists are not making their own jobs easier. This is why librarians must enlist the help of societies and associations such as STWP in reminding others, and remembering themselves, that each piece of documentation is not a unique bit of printing destined to exist forever in its own little cubicle. The present store of documentation is so vast that it is almost incomprehensible, but every new publication must be added to it after the first use. To provide recall from this mass, the major subject or subjects of individual documents must be indicated to the indexer, whether this indexer is a machine or a man.

Examination of the growth of the application of the key-word indexes reveals their importance and indicates the stress that must be placed upon them. In 1958, H. P. Luhn proposed the KWIC concept. Later in that year, Chemical Abstracts Service received a National Science Foundation grant to apply Luhn's concept to chemical information. By 1962, over 30 different permuted indexes were in use, and by 1965 the number had grown to several thousand. The current indexes include such products as "Chemical Titles," published by the American Chemical Society, "Index to Legal Theses & Research Projects," published by the American Bar Foundation, "Current Medical Terminology," published by the American Medical Association, and "Meteorological and Geostrophysical Titles," published by the American Meteorological Society. An announcement received recently by the Douglas Library listed permuted indexes available to technical publications such as the Military Electronics Conferences, National Telemetering Conferences, Institute of Radio Engineers, and Institute of Electrical and Electronic Engineers' Transactions, Environmental Sciences Proceedings, and others. Many more of these indexes have been developed for use in industry. Many are produced for internal use only, as, for instance, at Douglas. The extent of the influence of these indexes must be recognized, whether or not the form is accepted as a proper index.

Although the indexing requirement of a title should not take precedence over its main purpose--to describe the contents of the document--nor be considered the only important feature to an indexing system, its value in automatic indexing cannot be overstressed. The Abstract, the Table of Contents,

the Introduction, and the Conclusion also are relied upon to provide guidance for the indexer under almost all circumstances. However, under conventional cataloging conditions, the human indexer is able to exercise some judgment with regard to the relative merits of terms, descriptors, subjects, and so forth. The major emphasis here is directed towards the mechanically created index, where human judgment cannot be applied once the indexing program is established. So even those who question the value of permuted indexes should provide titles that can be used.

Another serious problem arises in relation to the publishing of all these papers and documents. Every author enjoys having his creation appear in print, and without qualification, "the more the better." Certainly, each publication is directed toward a different audience and justification for duplication is easy. But what does duplication do to the cataloger, the reference librarian, and the researcher? Imagine the frustration of finding five citations to a subject, only to discover that the five citations cover two articles published in five different sources! Such multiple publication of the same article causes problems for the scientist and engineer as well. Also, the engineer's sources or references may be credited to a corporate author or publication different from that used by the library for the same article.

Figure 7 shows an example of the many corporate authors possible for a single article. For this particular item, the work was done by a member of the staff of the California Institute of Technology, under funding from the Air Force Aerospace Research Labs, and submitted for publication to the "American Journal of Physics." The report was transmitted per contractual agreement to the Defense Documentation Center for dissemination throughout the defense community. Because of its application to aerospace endeavors, the National Aeronautics and Space Administration also acquired the document for dissemination. Thus, there are five sources established, three of which are primary and two (DDC and NASA) are secondary.

Are not all of these sources entitled to claim this document as their own? Certainly they are. The Aerospace Research Lab paid for the research, the California Institute of Technology prints its staff's output, the "American

PDL 56129

CALIFORNIA INST OF TECHNOLOGY JET PROPULSION LAB AEROSPACE RESEARCH LABS AMERICAN JOURNAL OF PHYSICS /DEFENSE DOCUMENTATION CTR/ /NATIONAL AERO AND SPACE ADM/	TR 32-738 ARL 65-151 AD 622 277 N65-31848
PLASMA NONUNIFORMITY AND GRID EROSION IN AN ELECTRON BOMBARD- MENT ION ENGINE. 7P.	PLASMA NONUNIFORM GRID EROSION ELECTRON ION ENGINE ELECTRON- -BOMBARDMENT EMISSION SURFACE DENSITY FLOW RATE
D. J. KERRISK, T. D. MASEK	
30 JULY 1965	
C.1, C.2	

Figure 7. Index Card with Multiple Corporate Authors

"Journal of Physics" agreed to publish the article for circulation among the scientific community. Both DDC and NASA are obligated under their charters to announce and disseminate all work performed through government funding.

Those responsible for the publication of material cannot solve this problem by limiting the number of times a report is published, but they can ensure that each publication makes note of the other sources in which it appears. When previous or prospective publications are known, why should not this information be included in a footnote? If the specific report number or issue number were not available, it would be helpful if a statement such as this were used: "Also published as Jet Propulsion Lab Technical Report and Aerospace Research Lab Technical Report," or "Announcement is anticipated in the Defense Documentation Center's Technical Abstract Bulletin and the National Aeronautics and Space Administration's Scientific and Technical Aerospace Reports Bulletin." Either of these statements would provide both cataloger and scientist with a warning that the document may appear under any one of those various corporate authors.

Often library users do not recognize the same material in different forms. Having once seen a journal article, a user may not be aware that the grey-cover ARL Report contains the same material--at least, not until he has requested the document and has begun re-reading it. This can also be true of the cataloger and librarian. Yet, with the tremendous quantity of information published today, information specialists must do all they can to save the reader from such circumstances. Libraries are given generous budgets, professional staffs, and expensive computer programs to inform their clients of material of interest to them. Informing an individual over and over again of the same material is not pleasing to either the individual or the responsible information specialist. The integrity of the entire information handling system is downgraded by these occurrences. To help avoid this, the author and the technical writer should give notice of the multiple publication of any documents he handles.

With the widespread use of mechanized systems has come another problem--the variations used in individual corporate names. Some mechanized programs have avoided the problem to some degree by coding the corporate authors. However, this imposes another set of problems. The concern of the information specialist is with the use of company names, agency names, universities, and so forth, as they are input to a computer. Take, for example, "IBM." In most communications, the use of "IBM" conveys the desired message. But the name of the company is International Business Machines, so corporate source notations should appear with the full name, "International Business Machines." This is a simplified example, because most catalogers would know that the citation should use the full name. It is not always so obvious, however, whether initials can constitute an acceptable corporate entry. IBM is not acceptable, because this is not the legal name of the organization. IIT Research Institute is acceptable and correct, since this organization officially discontinued the use of the title "Illinois Institute of Technology Research Center" and adopted ITT Research Institute as its official name. These examples illustrate how confusing to new and

inexperienced catalogers abbreviations in corporate names can be. Such abbreviations should be avoided unless they constitute a legal designation.

Even such variations as abbreviating "company" one time and spelling it out the next time can cause confusion. When a computer sorts, it sorts according to standard computer rules. Douglas could appear at least four different ways (Figure 8), without counting eliminating the periods or including divisional entities. The computer does not recognize all these names for the same company--it recognizes only a change in the characters in each.

Authors, technical writers, and information specialists must all be certain to use the corporate name in the same manner each time, so that the items will all appear together in a computer printout. It is important for indexers and catalogers to be precise about entries, but the tendency is to accept an

DOUGLAS AIRCRAFT COMPANY
DOUGLAS AIRCRAFT CO.
DOUGLAS AIRCRAFT COMPANY, INC.
DOUGLAS AIRCRAFT CO., INC.

Figure 8. Variations of Single Corporate Author

entry as it appears on the document. For this reason, the careful use of the corporate name can be of great help in facilitating the retrieval of valuable information.

Projections into the future indicate that, without doubt, the next major breakthrough in information systems will be total text input of some sort. Paper tapes and optical scanning methods have been tried with varying degrees of success but they reveal that there are several obvious questions to be answered in relation to full text input. These are as follows:

1. What about the extent to which the "stop list" would have to be enlarged beyond that of the KWIC stop list?
2. What about synonyms and near-synonyms? In the interest of variety, the writer will strive for a variation in terms, perhaps using words such as automobile, car, vehicle, sedan, all in reference to the same subject. When retrieving from total text input, care must be taken to ensure that each of these terms is used for recall.
3. What about antonyms? When a request arises for information on safety in handling chemicals, will there by a reference also to "hazards in handling chemicals"?
4. What about adjectives and adverbs that lose meaning when taken out of context? These words create problems, not so much in retrieval as in clutter (Figure 9). They are a vital part of all writing for logical communications, yet they are rarely used for retrieval. One interesting exception is the term "orbital." When first entered into the Douglas dictionary, the only term used was "orbital," although the usual procedure was to enter both the plural and singular forms of all terms. It seemed obvious that no one would ever use "orbital" by itself, much less in plural form. Yet within no more than 6 months, Douglas librarians were not only coping with "orbital paths," as expected, but also with "orbitals." The plural form has now become an accepted term in the aerospace industry.

KEY WORD LIST

1.16	DERIVE	HIGH	PROBABLY
1	DETERMINED	HIGH	PROBLEMS
29	DETERMINED	HIGH	PROCESSING
1	DETERMINING	HIGH	PROCESS
1	DETERMINE	HIGH	PROFILE
8	DETERMINING	HIGH	PROJECT
5	DETERMINED	HIGH	P
4	DETERMINING	HIGH	RATE
ACCORDING	DIFFICULTY	HIGH	REFERENCE 1
ADDITIONAL	DISTINCTION	HIGH	REFERENCE 2
ADDITION	DISTINCTIVE	HIGH	REFERENCE 3
AMERICAN	DOCUMENTATION	HIGH	REFERENCE
AMONG	DOCUMENTATION	HIGH	REFLECTING
APPEARED	DOCUMENTED	HIGH	REFLECTS
APPEARING	DOCUMENTED	HIGH	RELATIONSHIP
APPAREL	ELIMINATED	HIGH	RELATIVITY
APPLICATION	ELIMINATING	HIGH	RELEVANCE
ARRANGED	ENGINEER	HIGH	REPORTED
ARRANGE	EXAMPLE	HIGH	REPORT
ASSOCIATED	EXISTING	HIGH	REQUESTED
ASSOCIATION	EXPANDED	HIGH	REQUESTING
AUTOCODED	EXPERIENCE	HIGH	REQUESTS
AUTOCODING	EXPERIMENTATION	HIGH	REQUIRED
AUTOMATICALLY	EXPERIMENT	HIGH	REQUIRMENT
AUTOMATIC	EXPERIMENT	HIGH	RESULTS
BASED	FABRIC	HIGH	RESEARCH
BELIEVED	FACTOR	HIGH	RESULTANT
BELIEVE	FACT	HIGH	RESULT
8	FILE	HIGH	RETRIEVAL
CALLED	FIRST	HIGH	SEARCHED
CLOSELY	FOLLOWING	HIGH	SEARCHING
COLLECTION	FOLLOW	HIGH	SEARCH
COLLECT	FORM	HIGH	SECOND
COMPARED	FORMATA	HIGH	SELECTED
COMPARE	FOUND	HIGH	SELECTING
COMPARING	FREQUENCY	HIGH	SELECT
COMPUTATION	FREQUENTLY	HIGH	SEMANTICALLY
COMPUTER	FREQUENT	HIGH	SEMANTICAL
COMPUTING	FRiction	HIGH	SEMANTIC
CONSIDERED	F	HIGH	SINGLE
CONSIDERING	FURTHER	HIGH	SPECIFIC
CONTAINED	FUTURE	HIGH	SPECIFIED
CONTAINING	GENERAL	HIGH	SPECIFY
CONTAIN	GENERATED	HIGH	STATE
COORDINATE	GENERATE	HIGH	STATISTICALLY
COORDINATE	GENERATING	HIGH	
C	GENERATION	HIGH	
DEGREE	GIVEN	HIGH	
DERIVED	GROUP	HIGH	

Figure 9. Automatic Index

CONCLUSIONS

The technical writer's key to computerized information retrieval is not the restriction of all writing to a 10,000-word vocabulary. Rather, the key is the thoughtful selection and use of truly pertinent and descriptive terms, especially in those portions of a report which serve as key tools for effective storage and retrieval of valuable information, and in fully documented publication notations. Those who feverishly embrace the new word, or the new concept described by an old word, must consider the problems of the information retrieval specialists, and must realize that the kind of vocabulary used today must be retrievable tomorrow.

BIBLIOGRAPHY

- M. Fischer. The KWIC Index Concept: A Retrospective View. American Documentation, April 1966, Pages 57-70.
- A. G. Hoshovsky. How Technical Writers Can Use and Improve Technical Retrieval Systems. Office of Aerospace Research, 17 June 1966.
- M. E. Stevens. Automatic Indexing: A State-of-the-Art Report. National Bureau of Standards, NBS Monograph 91, 30 March 1965.
- Chemical Titles. American Chemical Society, No. 26, 19 December 1966.
- Second National Symposium on Engineering Information. Engineers Joint Council, March 1966.
- Symposium on Materials Information Retrieval. Aeronautical Systems Division, ASD-TDR-63-445, May 1963.